D1-82-0319

BOEING SCIENTIFIC RESEARCH LABORATORIES

# A "String Algorithm" for Shortest Paths in Directed Networks



Victor Klee

Mathematics Research

November 1963

D1-82-0319

# A "STRING ALGORITHM" FOR SHORTEST PATHS IN DIRECTED NETWORKS

by

Victor Klee
University of Washington

Mathematical Note No. 333

Mathematics Research Laboratory

BOEING SCIENTIFIC RESEARCH LABORATORIES

November 1963

The literature contains several algorithms for finding the shortest path between two nodes P and Q of a network, where the distances or arc-lengths are assumed to be positive. (For references, consult the review article by Pollack and Weibenson [3] and the book by Ford and Fulkerson [1].) Some of the algorithms, and in particular some of the analogue devices, are applicable only when the distance matrix is symmetric. As was remarked in [1] and [3], this is true of the simplest of the analogue procedures ----- the "string algorithm" reported by Minty [2]. It consists of making an inelastic string model of the network, with knots corresponding to nodes and string-lengths proportional to the corresponding distances, and then stretching the knots P and Q as far apart as is possible without breaking the string; this produces at least one straight path from P to Q, and each such straight path corresponds to a shortest path in the network.

Networks with asymmetric distance matrices are most conveniently represented by means of directed networks, in which every arc is regarded as a one-way street of the appropriate length. In the present note we describe a simple cutting procedure (related to one suggested by Thomas Seidman) which can be combined with any algorithm for undirected networks (symmetric distance matrix) so as to form a shortest-path algorithm for directed networks (asymmetric distance matrix). In particular, the cutting and stretching can be alternated to form a "string algorithm" for directed networks.

* * * * * * * * * * * * * * * * *

For each directed network $\underset{\sim}{N}$, let $\underset{\sim}{N}^u$ denote the corresponding undirected network.

THEOREM *Suppose that* $P$ *and* $Q$ *are nodes of a finite directed network* $\mathbf{N}_1$ *which has* $v$ *nodes and* $e$ *arcs, and that there is a path from* $P$ *to* $Q$ *in* $\mathbf{N}_1$. *Suppose that* $\mathbf{A}$ *is an algorithm for finding shortest paths in undirected networks, and let the sequential procedure* $\mathbf{S}_1, \mathbf{C}_1, \mathbf{S}_2, \mathbf{C}_2, \ldots$ *be as follows:*

$\mathbf{S}_i$) *Apply* $\mathbf{A}$ *to the undirected network* $\mathbf{N}_i^u$ *to find a shortest path* $\pi_i$ *from* $P$ *to* $Q$ *in* $\mathbf{N}_i^u$. *Suppose* $\pi_i$ *is given by*

$$V_0^i \alpha_1^i V_1^i \alpha_2^i V_2^i \ldots V_{h(i)-1}^i \alpha_{h(i)}^i V_{h(i)}^i,$$

*where the arcs* $\alpha_j^i$ *and the nodes* $V_j^i$ *are listed as they appear in traversing* $\pi_i$ *from* $P = V_0^i$ *to* $Q = V_{h(i)}^i$.

$\mathbf{C}_i$) *If* $\pi_i$ *is also a path in* $\mathbf{N}_i$, *terminate the procedure. If* $\pi_i$ *is not a path in* $\mathbf{N}_i$, *there exists a smallest index* $r(i)$ *and a largest index* $s(i)$ *(possibly the same) such that the directions of* $\alpha_{r(i)}^i$ *and* $\alpha_{s(i)}^i$ *in* $\pi_i$ *are opposite to their directions in* $\mathbf{N}_i$. *Let* $\mathbf{N}_{i+1}$ *be the directed network that is obtained from* $\mathbf{N}_i$ *by deleting every arc of* $\mathbf{N}_i$ *that (like* $\alpha_{r(i)}^i$*) ends in* $\mathbf{N}_i$ *at* $V_{r(i)-1}^i$ *but is not* $\alpha_{r(i)-1}^i$, *and deleting every arc of* $\mathbf{N}_i$ *that (like* $\alpha_{s(i)}^i$*) starts in* $\mathbf{N}_i$ *at* $V_{s(i)}^i$ *but is not* $\alpha_{s(i)+1}^i$. *There is a path* $\Sigma_i$ *from* $P$ *to* $Q$ *in* $\mathbf{N}_{i+1}$ *such that* $\Sigma_i$ *is actually a shortest path from* $P$ *to* $Q$ *in* $\mathbf{N}_i$.

*The procedure terminates at some stage* $\mathbf{C}_t$ *for which*

$$t \leq \min(v, (e+2)/2),$$

*and the path* $\pi_t$ *is a shortest path from* $P$ *to* $Q$ *in the directed network* $\mathbf{N}_i$.

(*The same conclusion holds if* $\mathbf{C}_i$ *requires only the first of the two deletions specified above, or if it requires only the second.*)

Proof. Of course the algorithm itself does not involve the actual *construction* of $\Sigma_i$, but we require the existence of $\Sigma_i$ (when $\mathbf{C}_i$ does not specify termination) to show that the sequential procedure $\mathbf{S}_1$, $\mathbf{C}_1$, $\mathbf{S}_2$, $\mathbf{C}_2$, ... can actually be followed and that each of the paths $\Sigma_1, \Sigma_2, \ldots$ is a shortest path from P to Q in $\mathbf{N}_1$. Since $\mathbf{N}_1$ is finite, the procedure must terminate at some stage $\mathbf{C}_t$ and then $\pi_t$ is a shortest path from P to Q in $\mathbf{N}_i$.

Suppose $\mathbf{C}_i$ does not specify termination and let $\Sigma_{i-1}$ be a shortest path from P to Q in $\mathbf{N}_i$, given by

$$W_0^i\ \beta_1^i\ W_1^i\ \beta_2^i\ W_2^i\ \ldots\ W_{l(i)-1}^i\ \beta_{l(i)}^i\ W_{l(i)}^i,$$

where of course $W_0^i = P$ and $W_{l(i)}^i = Q$. In constructing $\Sigma_i$, we consider the following three possibilities:

(i) no $W_j^i$ is equal to either $V_{r(i)-1}^i$ or $V_{s(i)}^i$;

(ii) there exists $j$ such that $W_j^i = V_{r(i)-1}^i$ and $j < k \Rightarrow W_k^i \neq V_{s(i)}^i$;

(iii) there exist $j$ and $k$ such that $j < k$, $W_j^i = V_{r(i)-1}^i$ and $W_k^i = V_{s(i)}^i$.

When (i) holds, we define $\Sigma_i = \Sigma_{i-1}$. When (ii) holds, we obtain $\Sigma_i$ by following $\pi_i$ from P to $V_{m(i)-1}^i$ and then following $\Sigma_{i-1}$ from $V_{m(i)-1}^i$ to Q. When (iii) holds, we obtain $\Sigma_i$ by following $\pi_i$ from P to $V_{m(i)-1}^i$, next following $\Sigma_{i-1}$ from $V_{m(i)-1}^i$ to $V_{n(i)}^i$, and then following

$\pi_i$ from $V^i_{n(i)}$ to $Q$. In each case, it is easily verified that $\Sigma_i$ has the stated properties. Thus the existence of $t$ is established and it remains only to show that $t \leq \min(v,(e+2)/2)$.

Let us review the special properties of certain nodes and arcs of $\underset{\sim}{N}_i$ relative to $\underset{\sim}{N}_i$ itself and relative to $\underset{\sim}{N}_j$ for $j > i$.

(a) $V^i_{r(i)-1} \neq Q$. If $V^i_{r(i)-1} = P$, then at least one arc of $\underset{\sim}{N}_i$ ends at $V^i_{r(i)-1}$ but no arc of $\underset{\sim}{N}_j$ ends there. If $V^i_{r(i)-1} \neq P$, then at least two arcs of $\underset{\sim}{N}_i$ end at $V^i_{r(i)-1}$ but at most one arc of $\underset{\sim}{N}_j$ endsthere.

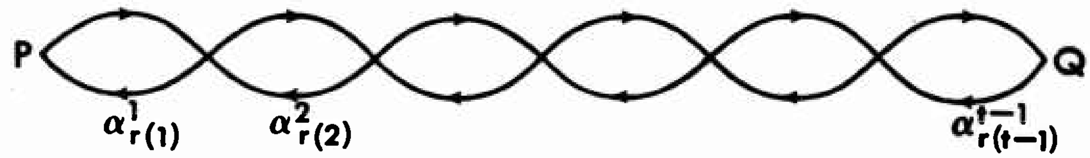
(b) $\alpha^i_{r(i)}$ ends at $P$ or is coterminal with another arc of $\underset{\sim}{N}_i$; $\alpha^i_{r(i)}$ does not end at $Q$ and does not appear in $\underset{\sim}{N}_j$.

(c) $\alpha^i_{r(i)-1}$ does not end at $P$ or $Q$, and is nonexistent if $\alpha^i_{r(i)}$ ends at $P$. If $\alpha^i_{r(i)}$ does not end at $P$, then $\alpha^i_{r(i)-1}$ is coterminal with another arc of $\underset{\sim}{N}_i$ but not with another arc of $\underset{\sim}{N}_j$.

We see from (a) that the $t$ nodes $Q, V^1_{r(1)-1}, V^2_{r(2)-1}, \ldots, V^{t-1}_{r(t-1)-1}$ are pairwise distinct, and consequently $t \geq v$. From (b) and (c) it follows that the arcs $\alpha^i_{r(1)-1}, \alpha^1_{r(1)}, \alpha^2_{r(2)-1}, \alpha^2_{r(2)}, \ldots, \alpha^{t-1}_{r(t-1)-1}, \alpha^{t-1}_{r(t-1)}$ are pairwise distinct. If $\alpha^i_{r(i)}$ ends at $P$, then $\alpha^i_{r(i)-1}$ does not appear, but this happens for at most one value of $i$, and since at least one arc of $\underset{\sim}{N}_1$ ends at $Q$ we conclude that $e \geq 2t - 2$.

The above reasoning completes the proof when $\underset{\sim}{C}_i$ is as originally described, and also when $\underset{\sim}{C}_i$ is replaced by $\underset{\sim}{C}_i'$ which requires only the first of the specified deletions. Similar reasoning applied to $\underset{\sim}{C}_i''$, which

Figure 1

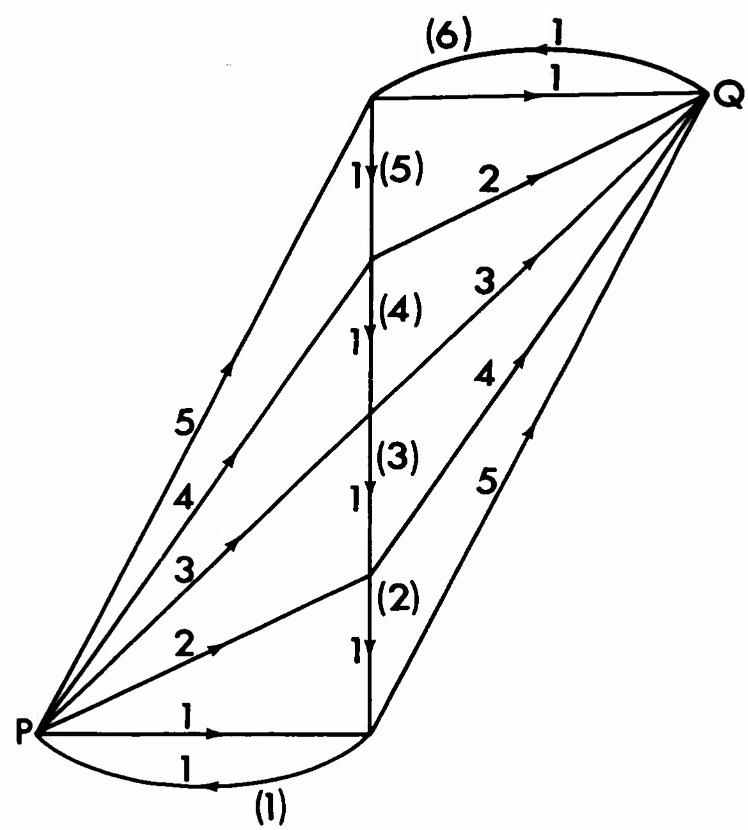



v nodes. 2(v-1) arcs, all of the same length.

$t = v = (e+2)/2$

($\pi_i$ follows upper arcs except at $\alpha^i_{r(i)}$)

Figure 2

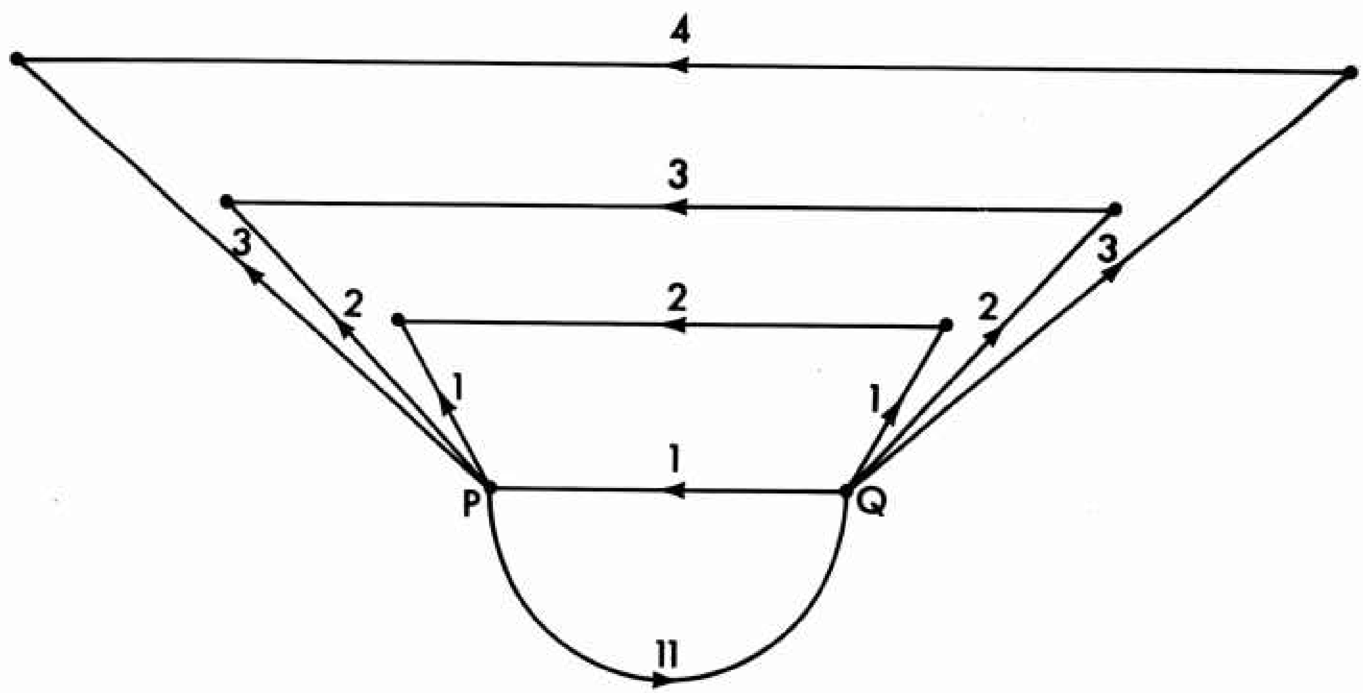



v nodes. 3v - 5 arcs, with lengths as indicated.

$\alpha^i_{r(i)}$(= $\alpha^i_{s(i)}$) indicated by (i)

$t = v$

Figure 3




2n nodes.

3n - 1 arcs.

$$t = n + 1 = \frac{v+2}{2} = \frac{e+4}{3}$$

($\pi_i$ is uniquely determined)

## REFERENCES

1. L. R. Ford, Jr. and D. R. Fulkerson, Flows in Networks, Princeton University Press, 1962.

2. G. J. Minty, "A Comment on the Shortest-Route Problem," Opns. Res. 5, 724 (1957).

3. M. Pollack and W. Wiebenson, "Solutions of the Shortest-Route Problem-- A Review," Opns. Res. 8, 224-230 (1959).